Cite this paper as: Ranjbarfard, M. & Parvaneh, E. (2018), "Applications of data mining in insurance", *7th Iranian Joint Congress on Fuzzy and Intelligent Systems, Bojnord, Iran*

Applications of data mining in insurance

Mina Ranjbarfard[1], Elahe Parvaneh

Department of management, Alzahra University, Tehran, Iran

Abstract:

Today, with the continuous advancement of technology, the insurance industry is facing an increase in the volume and complexity of insurance data in its various fields. Analyzing this vast amount of data produces interesting results that can be very important for the survival of insurance companies; Therefore, insurance companies need to use technologies based on which they can extract, transmit and convert data into information. One of the types of information technologies that can be used in insurance services is a platform of business intelligence using data mining techniques. In this paper, we tried to review the applied research conducted in the insurance field as a guide for researchers and managers in this field. Accordingly, after searching for insurance articles with data mining approach in reputable scientific databases, about 28 articles were extracted and studied as research background. These articles are categorized based on their respective applications and provided in this research. As a result, 8 applied areas were obtained regarding the use of data mining methods in insurance.

Keywords: data mining, insurance industry, the role of information technology, knowledge discovery

[1] m.ranjbarfard@alzahra.ac.ir

کاربردهای داده‌کاوی در بیمه

مینا رنجبر فرد[۱]، دانشگاه الزهرا
الهه پروانه‌وار، دانشگاه الزهرا

چکیده: امروزه با پیشرفت مستمر فنّاوری، صنعت بیمه با افزایش حجم و پیچیدگی داده‌های بیمه‌ای در حوزه‌های مختلف آن روبرو شده است. تحلیل این حجم وسیع داده‌ها نتایج جالب‌توجهی ایجاد می‌کند که می‌تواند برای بقای شرکت‌های بیمه اهمیت زیادی داشته باشد؛ بنابراین شرکت‌های بیمه‌ای نیازمند استفاده از فناوری‌هایی هستند که بر مبنای آن بتوانند داده‌ها را استخراج کنند، انتقال دهند و تبدیل به اطلاعات نمایند. یکی از انواع فناوری‌های اطلاعات قابل‌استفاده در خدمات بیمه‌ای، بستری از هوش تجاری با استفاده از فنون داده‌کاوی است. در این مقاله سعی شده است مروری بر پژوهش‌های کاربردی انجام‌شده در حوزه بیمه صورت گیرد که راهنمایی برای محققان و مدیران این حوزه باشد. بر همین اساس پس از جست‌وجوی مقالات بیمه‌ای با رویکرد داده‌کاوی در پایگاه‌های علمی معتبر حدود ۲۸ مقاله به‌عنوان پیشینه پژوهش استخراج و مطالعه شد. این مقالات بر مبنای کاربردهای مربوطه دسته‌بندی و طی این پژوهش ارائه می‌شود. درنتیجه ۸ حوزه کاربردی در خصوص استفاده از روش‌های داده‌کاوی در بیمه به دست آمد.

واژه های کلیدی: داده‌کاوی، صنعت بیمه، نقش فناوری اطلاعات، کشف دانش

[۱] استادیار گروه مدیریت، دانشگاه الزهرا، عضو هیئت علمی دانشگاه الزهرا

mina.ranjbar.ie@gmail.com

مقدمه

صنعت بیمه به‌طورکلی به توانایی تبدیل داده خام به هوشمندی (هوشمندی در ارتباط با مشتری، بازار، رقبا و محیط کسب‌وکار) بستگی دارد. طی گذر زمان فنّاوری پردازش داده به‌طور فوق‌العاده‌ای پیشرفت کرده است و ابزارهایی مانند انبار داده، OLAP و داده‌کاوی که زیربنا و شالوده یک محیط هوشمند مؤثر کسب‌وکار را شکل می‌دهند، به‌صورت گسترده‌ای موردقبول صنعت قرارگرفته‌اند. بااین‌حال، شرکت‌های بیمه در سازگاری با این ابزارها نسبتاً کند بوده‌اند که عمدتاً به علت عدم رقابت ناشی از مقررات محافظتی است؛ اما اکنون، آن‌ها دیگر نمی‌توانند این موضوع را نادیده بگیرند که اینترنت، تثبیت و همگرایی بیمه با سایر خدمات مالی، به‌سرعت ساختار اصلی صنعت را تغییر می‌دهد. بیمه‌گذاران می‌توانند از منابع داخلی داده در بهبود، ساده و مؤثر کردن و نیز در جهت اهداف فرایند کسب‌وکار در مواردی از قبیل قیمت‌گذاری بیمه ثابت و واقعی، کاهش هزینه‌های مؤثر تحویل کالاهای بیمه و نیز هزینه‌های کلی، تسهیل پیگیری موارد تقلب و ازاین‌دست موضوعات که قبلاً غیرقابل دستیابی بود، استفاده کنند. در این راستا هوش تجاری[2] می‌تواند به‌صورت گسترده‌ای به‌ویژه توسط مدیران اجرایی سازمان‌ها برای توصیف برنامه‌های کاربردی تحلیلی مورداستفاده قرار گیرد. BI تبدیل به استراتژی اولیه سازمان‌ها شده و همچنین به‌عنوان یک ابزار در اداره مؤثر و خلاقانه برای مدیران اجرایی و رهبران کسب‌وکار شناخته‌شده است (Watson & Wixom, 2007). یکی از ابزارها یا به عبارت بهتر هسته مرکزی هوش تجاری، داده‌کاوی است. داده‌کاوی اصطلاحی است که به استخراج اطلاعات از میان انبوهی از داده اطلاق می‌شود. داده‌کاوی می‌تواند برای ارزیابی و اصلاح قواعد و فرآیندهای سازمان که اساس خط‌مشی پذیره‌های بیمه آن‌ها است، به کار برود. مزیت متمایز داده‌کاوی قابلیت آن در آشکار ساختن الگوها و روابط در حجم بالای داده‌ها و ایجاد پیش‌بینی کننده‌های مختلف با سطوح اطمینان خاص است که مربوط به وظایف عملکردی بیمه‌گری در کسب‌وکار بیمه است (Sumathi & Sivanandam, 2006).

[2] Business Intelligence (BI)

در رویکرد پژوهشی پیش رو از یافته‌های پژوهش‌های قبلی برای بهبود وضع موجود استفاده می‌شود. در این مقاله تلاش بر آن بوده که مطالعات برحسب کاربردهای داده‌کاوی توصیف و در اختیار خواننده قرار بگیرد. ازاین‌رو پس از توصیف و توسعه مفاهیم اصلی داده‌کاوی، پژوهش‌های مطالعه شده بر اساس کاربردهای داده‌کاوی در بیمه ارائه‌شده‌اند. به‌منظور ساده‌سازی و استفاده از قوانین و فنون مربوط به داده‌کاوی به نکات اصلی و ویژه هر پژوهش اشاره می‌شود. در ادامه سعی بر آن شده است که ادبیات موجود در پایگاه‌های علمی گوناگون از قبیل Scopus، Emerald، Science Direct، IEEE، Google Scholar موردبررسی قرارگرفته و پژوهش‌های مرتبط بیمه با رویکرد داده‌کاوی استخراج شود.

داده‌کاوی

داده‌کاوی مجموعه‌ای از الگوریتم‌ها است که به سازمان‌ها کمک می‌کند تا بر مهم‌ترین اطلاعات و دانش موجود در پایگاه‌های داده خود تمرکز کنند. جکسون در تعریف داده‌کاوی دو روش متفاوت در انجام داده‌کاوی بیان می‌کند که طی آن یا به دنبال ساخت مدل و یا یافتن الگو از میان داده‌ها است. در رویکرد اول استخراج مدل از میان مجموعه داده‌ها برای شناسایی و توصیف ویژگی‌های اصلی توزیع داده‌ها انجام می‌شود. به‌عبارت‌دیگر به دنبال نشان دادن روابط است. دومین رویکرد داده‌کاوی، برای شناسایی الگوهای غیرمعمول رفتار از میان قوانین به کار می‌رود (Jackson, 2002). داده‌کاوی در سال‌های اخیر به‌عنوان یک رویکرد بسیار قدرتمند برای استخراج اطلاعات معنی‌دار از پایگاه داده‌های بزرگ و انبار داده‌ها پدید آمده است. داده‌کاوی دارای سه هدف اصلی توصیف، پیش‌بینی و تجویز است (Amani & Fadlalla, 2017). تکنیک‌های داده‌کاوی به‌عنوان زیرمجموعه‌ای از این اهداف شامل چند دسته کلی می‌شوند. هان عملگرهای داده‌کاوی را شامل مواردی از قبیل شخصی‌سازی و تمایز[3]؛ استخراج الگوهای تکراری، باهم‌آیی‌ها و همبستگی‌ها[4]؛ طبقه‌بندی و رگرسیون[5]؛ تجزیه‌وتحلیل خوشه‌ای[6]؛ و

[3] Characterization and discrimination

[4] Fequent patterns, associations, and correlations

[5] Classification and regression

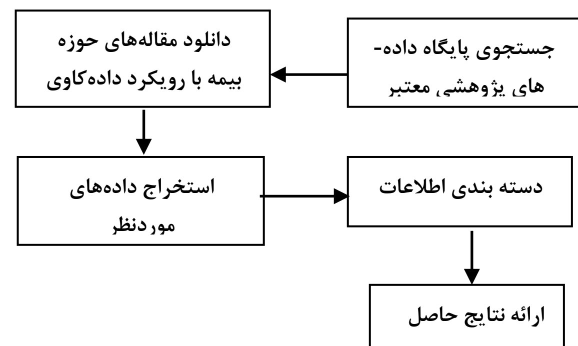
[6] Clustering analysis

تجزیه تحلیل نقاط پرت[7] می‌داند (Han, Pei, & Kamber, 2011).

روش تحقیق

پژوهش حاضر به دنبال بررسی مطالعات موجود در حوزه صنعت بیمه با رویکرد داده‌کاوی است، به‌گونه‌ای که بتواند جنبه‌های گوناگون قابل‌بررسی در صنعت بیمه را ارائه دهد؛ بنابراین می‌توان گفت رویکرد این پژوهش کاربردی توصیفی است. این پژوهش کاربردی است زیرا توسعه کاربرد دانش در یک حوزه خاص را بیان می‌کند. همچنین توصیفی است به این لحاظ که توصیف نتایج پژوهش صرفاً به بازنمود مسئله تحقیق می‌پردازد و به دنبال تحلیل محتوای مقاله‌های یافت شده است.

در این پژوهش، تعدادی از مقاله‌های موجود در صنعت بیمه با رویکرد داده‌کاوی مطالعه شده است. پس از مرور نتایج حاصل از این مقالات، دسته‌بندی‌شده و اطلاعات آنان ارائه می‌شود. درنهایت روند انجام پروژه طبق فلوچارت زیر رسم است.



کاربرد داده‌کاوی در بیمه

به‌طورکلی اعتقاد بر این است که از نظریه‌پردازی تا عمل فاصله زیادی وجود دارد. تبدیل تئوری به عمل چنانچه انتظار می‌رود کار آسانی نیست. آن مستلزم همکاری مستمر متخصصان و محققان حوزه موردبررسی است. روند بررسی و شناخت بیمه موضوعی است که باید به‌دقت موردتوجه قرار بگیرد. از سوی دیگر، هدف نهایی برای فنّاوری داده‌کاوی خود فنّاوری نیست بلکه منظور ارتقاء کیفیت مدیریت، به‌ویژه کیفیت تصمیم‌گیری است. معمولاً یک شرکت بیمه می‌تواند محصولات یا خدمات متنوعی را ارائه دهد که به‌طور مداوم تغییر می‌کنند تا به ترجیحات و الزامات مختلف مشتریان برسند، به‌ویژه هنگامی‌که هر یک از استراتژی‌های بازاریابی به‌طور فزاینده‌ای مورداستفاده قرار می‌گیرند. اساساً یک شرکت بیمه باید از توانایی یادگیری در داده‌های خود، مزایای بالقوه آن و نحوه توزیع آن به‌درستی آگاهی داشته باشد تا بتوانند برنامه داده‌کاوی را به‌راحتی پیاده‌سازی کنند. داده‌کاوی در تصمیم‌های حیاتی کسب و کار به شرکت‌های بیمه کمک می‌کند و دانش تازه به‌دست آمده را موارد قابل پیاده‌سازی در کسب‌وکار مانند توسعه محصول بیمه‌ای، بازاریابی، مدیریت دارایی-بدهی وغیره تبدیل کند. ازاین‌رو مستلزم مطالعه در حوزه‌های مختلف بیمه‌ای است. این پژوهش سعی بر آن دارد که به مسائل مختلف بیمه‌ای بپردازد؛ بنابراین پس از مطالعه ۲۸ مقاله‌ی جمع‌آوری‌شده، ۸ حوزه تحقیقاتی در صنعت بیمه به دست آمد که در ذیل توضیحات مربوط به آن‌ها در اختیار قرار می‌گیرد.

قیمت‌گذاری. ویژگی اصلی یک اقدام بیمه‌ای این است که حق بیمه را در ابتدای قرارداد بیمه تعیین کند. برای تعیین حق بیمه دقیق سالیانه در شرکت بیمه، تخمین دقیق و قابل‌اعتماد از تعداد ادعاهایشان و میزان کل ادعا بسیار مهم است. کشف دانش بر مبنای داده‌های بیمه‌ای از طریق تکنیک‌های داده‌کاوی ازجمله خوشه‌بندی، طبقه‌بندی و رگرسیون در پژوهش‌های موردمطالعه انجام‌شده است (Yeo et al., 2002 ؛ Kahane, Levin, Meiri, & Zahavi, 2007 ؛ Sakthivel & Rajitha, 2017).

حفظ مشتری. یکی از مسائلی که در بیمه از اهمیت بسیاری برخوردار است، دستیابی به رشد بازار و سودآوری است. تنظیم سطح قیمت متناسب با هزینه‌ها و هم‌زمان با آن حفظ مشتری موجود و جذب مشتریان جدید مسئله‌ای است که در صنعت بیمه همچون دیگر صنایع باید به آن توجه شود. در جهت حفظ مشتریان بر اساس اطلاعات آن‌ها تکنیک‌های خوشه‌بندی، طبقه‌بندی، با هم‌آیی و همبستگی مورداستفاده قرارگرفته است (Smith et al., 2000 ؛ Kulkarni, R. V., & Devale, 2012 ؛ Song & Qiuhong, 2018).

---

[7] Outlier analysis

ارزیابی ارزش مشتری. برای تعیین قوانین و عواملی که ممکن است بر ارزش‌های مشتریان تأثیر بگذارد، تکنیک‌های متفاوت داده‌کاوی قابل‌بررسی خواهد بود. پژوهشگران مدل‌سازی داده‌های کسب‌وکار بیمه توسط الگوریتم‌های مختلف بیمه را پیشنهاد کرده‌اند. تکنیک‌های مورداستفاده از قبیل طبقه‌بندی، رگرسیون، خوشه‌بندی است (K.-S. Wu, Wang, & Shyng, 2013 ؛ K.-S. Wu et al., 2013 ؛ Fang, Jiang, & Song, 2016).

ارزیابی ریسک. هنگامی‌که یک مشتری جدید به یک شرکت بیمه برای یک محصول بیمه‌ای مانند اتومبیل جدید خود می‌آید یا زمانی که قصد تجدید قرارداد بیمه‌ای محصول موجود را دارد، شرکت بیمه باید بتواند احتمال ادعای خسارت مشتری را بسنجد؛ بنابراین سنجش احتمال ریسک مشتریان مطرح می‌شود. در این جهت پژوهش‌هایی انجام‌شده است که از تکنیک‌های مختلف داده‌کاوی یعنی طبقه‌بندی، رگرسیون، خوشه‌بندی استفاده‌شده است (Baecke & Bocca, 2017؛ Kašćelan, Kašćelan, & Burić, 2016؛ Ismail & Jemain, 2008).

افزایش مشتریان هدف. یکی از استراتژی‌های بسیار مهم بازاریابی که عموماً توسط شرکت‌ها برای جذب مشتریان مورداستفاده قرار می‌گیرد، بازاریابی شخصی است. داده‌کاوی ابزاری است که با استفاده از روش‌های مختلف انجام تفسیر، بهره‌برداری و ارزیابی دانش، بسیار مفید واقع می‌شود. تکنیک‌های مورداستفاده در این حوزه شامل طبقه‌بندی، خوشه‌بندی و قواعد باهم آیی است (C. H. Wu, Kao, Su, & Wu, 2005؛ Kim & Street, 2004؛ Liao, Chen, & Tseng, 2009؛ Liu & Cong, 2015).

تشخیص بیماری. استخراج الگو از پایگاه داده بیمه یک چارچوب جدید از داده‌کاوی ارائه می‌دهد که بر مبنای آن می‌توان حوادث محل کار و میزان آسیب‌ها و عوارض ناشی از آن را پیش‌بینی کرد. از سوی دیگر استفاده از پایگاه‌های داده بیمه در تشخیص بیماری و حتی تشخیص دوگانگی بیماری نیز کاربرد دارد. تکنیک‌های استخراج الگو در این حوزه، خوشه‌بندی، طبقه‌بندی، رگرسیون، قواعد انجمن و تحلیل بقا است (Nenonen, 2013؛ Kuo, Lin, & Shih, 2007؛ Chae et al., 2001 ؛ Silva & Jacinto, 2012).

تشخیص تقلب. کشف تقلب در بیمه مسئله‌ای است که همواره شرکت‌های بیمه با آن درگیر هستند. این امر در هر دو بعد فنی و عملی حائز اهمیت است. ازآنجایی‌که کشف تقلب به‌صورت دستی بسیار هزینه‌بر و زمان‌بر است و از سوی دیگر رشد داده‌ها بسیار سریع است، روش‌های تحلیلی داده‌کاوی می‌تواند مؤثر و مفید باشد. تکنیک‌های مورداستفاده در پژوهش‌های کشف تقلب عبارت‌اند از: خوشه‌بندی، طبقه‌بندی، رگرسیون، تحلیل داده‌های پرت، تحلیل شبکه اجتماعی (Nian, Zhang, Tayal, Coleman, & Li, 2016؛ Wang, Pai, Wu, Wu, & Li, 2017؛ Bayerstadler, van Dijk, & Winter, 2016؛ Shin, Park, Lee, & Jhee, 2012).

تعیین بیمه‌نامه. تعیین بهترین سیاست بیمه‌نامه برای خریداران بیمه بالقوه یک چالش بزرگ برای مشاوران بیمه است. مشتریان برای خرید بیمه‌نامه با زمینه‌های مختلف، شرایط اقتصادی، درجه اولویت ریسک و نیازهای خود می‌آیند. تکنیک‌های مورداستفاده در پژوهش‌های موردبررسی طبقه‌بندی و خوشه‌بندی است (Silva & Jacinto, 2012؛ Lin, Huang, & Lin, 2008).

عملکرد بیمه. شرکت‌های بیمه نیازمند بررسی کارایی در محیط فعالیت خود هستند. این امر به شرکت‌ها کمک می‌کند که بتوانند از نقاط قوت و ضعف خود به‌ویژه در محیط‌های رقابتی اطلاع یابند. یکی از روش‌های موفق در بررسی عملکرد شرکت‌های بیمه‌ای استفاده از تکنیک‌های داده‌کاوی است. تکنیک‌هایی که در پژوهش‌های موردمطالعه مشاهده شد عبارت‌اند از: طبقه‌بندی، رگرسیون و خوشه‌بندی (Wanke & Barros, 2016؛ Ennsfellner, Lewis, & Anderson, 2004).

نتایج

در رویکرد پژوهشی از یافته‌های پژوهش‌های قبلی برای بهبود وضع موجود استفاده نمود. در این مقاله تلاش بر آن بوده که مطالعات برحسب کاربردهای داده‌کاوی توصیف و در اختیار خواننده قرار بگیرد. ازاین‌رو پس از توصیف و توسعه مفاهیم اصلی داده‌کاوی، پژوهش‌های مطالعه شده بر اساس کاربردهای داده‌کاوی در بیمه ارائه‌شده‌اند. به‌منظور ساده‌سازی و استفاده از قوانین و فنون مربوط به داده‌کاوی به نکات اصلی و ویژه هر پژوهش مورد بررسی قرار

گرفت. در ادامه پایگاه‌های علمی گوناگون از قبیل Scopus، Emerald، Science Direct، IEEE، Google Scholar به جهت یافتن مقالات معتبر موردبررسی قرارگرفته و پژوهش‌های مرتبط بیمه با رویکرد داده‌کاوی استخراج شد.

این تحقیق شواهدی ارائه می‌کند که بر اساس آن می‌توان الگوریتم‌های داده‌کاوی موردنیاز برای هر بخش را شناسایی نمود. این امر فرایند بررسی بخش‌های مختلف بیمه را برای پژوهشگران ساده می‌کند. تکنیک‌های مختلف طبقه‌بندی، رگرسیون، خوشه‌بندی، باهم آیی و همبستگی، تحلیل داده‌های پرت، تحلیل شبکه اجتماعی و تحلیل بقا در حوزه‌های مختلف بیمه موردتوجه بوده است. نتایج حاصل در زیر بیان می‌شود.

- از بین تمامی این تکنیک‌ها سه تکنیک «طبقه‌بندی»، «رگرسیون»، «خوشه‌بندی» تقریباً در تمامی بخش‌ها مشترکاً مورداستفاده بوده‌اند.
- در خصوص الگوریتم‌های موردبررسی در این مقالات نیز انواع مختلف الگوریتم‌ها مورداستفاده هستند اما در بین تمامی الگوریتم‌ها دو نوع «درخت تصمیم» و «انواع رگرسیون» بیشترین کاربرد را داشته‌اند.
- بر اساس پژوهش‌های موردمطالعه دو بخش در بین تمامی حوزه‌های صنعت بیمه بیشتر موردتوجه بوده‌اند. حوزه «کشف تقلب» و «تشخیص بیماری» در بین تمامی مراجع مطالعات بیشترین تحقیقات را داشته‌اند.
- همچنین محصولات موردتوجه پژوهشگران طی جدول نشان داده‌شده است. چنانچه مشاهده می‌شود در بین انواع محصولات (رشته‌های) بیمه‌ای سه دسته بیشتر موردبررسی قرارگرفته است. این‌ها شامل «بیمه عمر»، «بیمه سلامت» و «بیمه خودرو» است.

داده‌کاوی می‌تواند برای ارزیابی و اصلاح قواعد و فرآیندهای سازمان که اساس خط‌مشی پذیره‌های بیمه آن‌ها است، به کار برود؛ بنابراین تصمیم‌گیرندگان سازمان می‌توانند با استفاده از نتایج حاصل از رویکرد مورداستفاده تصمیم‌های مؤثرتر و دقیق‌تری اتخاذ کنند. اساساً یک شرکت بیمه باید از توانایی یادگیری در داده‌های خود، مزایای بالقوه آن و نحوه توزیع آن به‌درستی آگاهی داشته باشد تا بتوانند برنامه داده‌کاوی را به‌راحتی پیاده‌سازی کنند. ازاین‌رو مستلزم مطالعه در حوزه‌های مختلف بیمه‌ای است. این مقاله حدود ۲۸ پژوهش در مورد کاربرد تکنیک‌های داده‌کاوی در صنعت بیمه را موردبررسی قرار داده است که در نتیجه آن پس از مطالعه ۲۸ مقاله‌ی جمع‌آوری‌شده، ۸ حوزه تحقیقاتی در صنعت بیمه از جمله قیمت‌گذاری، حفظ مشتری، ارزیابی ارزش مشتری، ارزیابی ریسک، افزایش مشتریان هدف، تشخیص بیماری، تشخیص تقلب، تعیین بیمه‌نامه و عملکرد بیمه به دست آمد.

جدول ۱ الگوریتم ها و تکنیک ها

نویسندگان	محصول بیمه‌ای	الگوریتم‌ها	تکنیک	کاربرد
Baecke & Bocca, 2017	خودرو	شبکه عصبی، رگرسیون، جنگل تصادفی	طبقه‌بندی، رگرسیون، خوشه‌بندی	ارزیابی ریسک
V.Kašćelan, L.Kašćelan & Burić, 2016	خودرو	KLR and SVR		
Ismail & Jemaim, 2008	خودرو	مدل‌های رگرسیون Lognormal، و Normal Gamma		
Wu, Kao, Su & Wu, 2005	بیمه عمر سالانه، بیمه سلامت، بیمه زنان	قواعد تصمیم‌گیری و الگوریتم ID3	طبقه‌بندی، خوشه‌بندی، قواعد باهم آیی	افزایش مشتریان هدف
Kim & Street, 2004	خودرو تفریحی	ترکیب الگوریتم‌های ژنتیک		

نویسندگان	محصول بیمه‌ای	الگوریتم‌ها	تکنیک	کاربرد
		GAs و شبکه‌های عصبی مصنوعی ANNs		
Liao, Chen & Tseng, 2009	بیمه عمر	K-means و قواعد با هم‌آیی		
Liu & Cong, 2015	بعضی انواع بیمه: اتومبیل، مسافرت، خانه، حوادث، شخصی، حمل بار	الگوریتم بهینه‌سازی اپریوری مبتنی بر آرایه Array		
Nenonen, 2013	بیمه سلامت	درخت تصمیم و قواعد باهم آیی	خوشه‌بندی، طبقه‌بندی، رگرسیون، قواعد انجمن، تحلیل بقا	تشخیص بیماری
Kuo, Lin & Shih, 2017	بیمه سلامت	خوشه‌بندی مبتنی بر مورچه (ASCA)، kmeans مورچگان (AK)، قواعد انجمن مبتنی بر کلونی		

نویسندگان	محصول بیمه‌ای	الگوریتم‌ها	تکنیک	کاربرد
		مورچگان		
Chae, Moon, Ho, et al., 2001	بیمه سلامت	رگرسیون لجستیک، درخت تصمیم شامل: CHAID، C5.0 و قوانین انجمنی		
Silva & Jacinto, 2012	بیمه سلامت	chi-square		
Nian, Zhang, Tayal et al., 2016	خودرو	ماشین بردار پشتیبان یک کلاسه، فاکتورهای پرت محلی (LOF)، کرنال شباهت، SRA	خوشه‌بندی، طبقه‌بندی، رگرسیون، تحلیل داده‌های پرت، تحلیل شبکه اجتماعی	تشخیص تقلب
Wang, Pai, Wu, et al., 2017	دندانپزشکی	ترکیب روش‌های سنتی و شبکه اجتماعی		
Bayerstadler, Dijk & Winter, 2016	بیمه سلامت	الگوریتم بیزین، الگوریتم		

نویسندگان	محصول بیمه‌ای	الگوریتم‌ها	تکنیک	کاربرد
		زنجیره مارکوف مونت‌کارلو		
Shin, Park, Lee, et al., 2012	بیمه سلامت	CDA scoreو درخت تصمیم		
Jandaghi Moradpour, 2015	بیمه عمر	خوشه‌بندی فازی نویز (FKM)	طبقه‌بندی، خوشه‌بندی	تعیین بیمه‌نامه
Lin, Huang & Lin, 2008	شامل عمر، مستمری، سلامت، حادثه و بیمه سرمایه‌گذاری	از شبکه‌های عصبی پیش‌خور با الگوریتم انتشار روبه عقب		
Yeo et al., 2002	خودرو	k-means، شبکه عصبی، برنامه‌ریزی عدد صحیح غیرخطی	خوشه‌بندی، طبقه‌بندی، رگرسیون	تعیین حق بیمه
Kahane, Levin, Meiri, et al., 2007	خودرو	مدل تحلیل بقا و یک مدل رگرسیون خطی		
Sakthivel & Rajitha,	بیمه غیرزندگی	شبکه عصبی		

نویسندگان	محصول بیمه‌ای	الگوریتم‌ها	تکنیک	کاربرد
2017	عمومی	مصنوعی (ANN) و شبکه عصبی پیش‌خور		
Smith, Willis & Brooks, 2000	خودرو	الگوریتم شبکه عصبی پیش‌خور سه لایه و نیز الگوریتم K-means با حداقل معیارهای مربع	خوشه‌بندی، طبقه‌بندی، با هم‌آیی، همبستگی	حفظ مشتری
Kulkami & Devale, 2012	بیمه عمر	قوانین انجمنی، خوشه‌بندی، طبقه‌بندی و همبستگی مناسب		
Song & Qiuhong, 2018	خودرو	درخت تصمیم و رگرسیون لجستیک		
Wank & Barros, 2016	انواع محصولات و مناطق ارائه خدمت	درخت رگرسیون (CART)، Bagging، جنگل تصادفی	طبقه‌بندی، رگرسیون، خوشه‌بندی،	عملکرد بیمه

نویسندگان	محصول بیمه‌ای	الگوریتم‌ها	تکنیک	کاربرد
		مبتنی بر CART، درخت‌های استنتاج مشروط (CIT)، جنگل تصادفی مبتنی بر CTI، شبکه عصبی مصنوعی، ماشین بردار پشتیبان، K نزدیک‌ترین همسایه، رگرسیون لجستیک		
Ennsfellner, Lewis & Anderson, 2004	بیمه‌های زندگی / سلامتی و غیرزندگی	مدل بیزی مرزی تصادفی		

منابع:

1. Amani, F. A., & Fadlalla, A. M. (2017). Data mining applications in accounting: A review of the literature and organizing framework. *International Journal of Accounting Information Systems*, *24*, 32–58. https://doi.org/10.1016/j.accinf.2016.12.004

2. Baecke, P., & Bocca, L. (2017). The value of vehicle telematics data in insurance risk selection processes. *Decision Support Systems*, *98*, 69–79. https://doi.org/10.1016/j.dss.2017.04.009

3. Bayerstadler, A., van Dijk, L., & Winter, F. (2016). Bayesian multinomial latent variable modeling for fraud and abuse detection in health insurance. *Insurance: Mathematics and Economics*, *71*, 244–252. https://doi.org/10.1016/j.insmatheco.2016.09.013

4. Chae, Y. M. Y.., Ho, S.. S. H., Cho, K. W., Lee, D. H., Ji, S. H., Laudon, K.., … Ohrr, H.. (2001). Data mining approach to policy analysis in a health insurance domain. *International Journal of Medical Informatics*, *62*, 103–111. https://doi.org/10.1016/S1386-5056(01)00154-X

5. Ennsfellner, K. C., Lewis, D., & Anderson, R. I. (2004). Production Efficiency in the Austrian Insurance Industry: A Bayesian Examination. *Journal of Risk and Insurance*, *71*(1), 135–159. https://doi.org/10.1111/j.0022-4367.2004.00082.x

6. Fang, K., Jiang, Y., & Song, M. (2016). Customer profitability forecasting using Big Data analytics: A case study of the insurance industry. *Computers & Industrial Engineering*, *101*, 554–564. https://doi.org/10.1016/j.cie.2016.09.011

7. Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Retrieved from https://books.google.de/books?hl=en&lr=&id=pQws07tdpjoC&oi=fnd&pg=PP1&dq=Data+Mining,+concepts+and+techniques&ots=tzGz3ZpDVZ&sig=5UpK01-16nfIEDKA3RgyW_tGH7c

8. Ismail, N., & Jemain, A. A. (2008). Construction of insurance scoring system using regression models. *Sains Malaysiana*, *37*(4), 413–420. Retrieved from http://digitalcommons.wayne.edu/jmasm/vol7/iss2/25/

Factor Analysis Technique. *WSEAS Transactions on Information Science & Applications*, *5*(10), 1415–1425. Retrieved from https://pdfs.semanticscholar.org/e212/96f1c7cca2401d54fd0018daec87a536c4ae.pdf

17. Liu, J., & Cong, M. (2015). The optimization of apriori algorithm based on array and its application in the analysis of insurance clients. In *2015 4th International Conference on e-Technologies and Networks for Development, ICeND 2015* (pp. 58–61). IEEE. https://doi.org/10.1109/ICeND.2015.7328539

18. Nenonen, N. (2013). Analysing factors related to slipping, stumbling, and falling accidents at work: Application of data mining methods to Finnish occupational accidents and diseases statistics database. *Applied Ergonomics*, *44*(2), 215–224. https://doi.org/10.1016/j.apergo.2012.07.001

19. Nian, K., Zhang, H., Tayal, A., Coleman, T., & Li, Y. (2016). Auto insurance fraud detection using unsupervised spectral ranking for anomaly. *The Journal of Finance and Data Science*, *2*(1), 58–75. https://doi.org/10.1016/j.jfds.2016.03.001

20. Sakthivel, K. M., & Rajitha, C. S. (2017). Artificial Intelligence for Estimation of Future Claim Frequency in Non-Life Insurance. *Global Journal of Pure and Applied Mathematics*, *13*(6), 973–1768. Retrieved from http://www.ripublication.com

21. Shin, H., Park, H., Lee, J., & Jhee, W. C. (2012). A scoring model to detect abusive billing patterns in health insurance claims. *Expert Systems with Applications*, *39*(8), 7441–7450. https://doi.org/10.1016/j.eswa.2012.01.105

22. Silva, J. F., & Jacinto, C. (2012). Finding occupational accident patterns in the extractive industry using a systematic data mining approach. *Reliability Engineering and System Safety*, *108*, 108–122. https://doi.org/10.1016/j.ress.2012.07.001

23. Smith, K. A., Willis, R. J., & Brooks, M. (2000). An analysis of customer retention and

9. Jackson, J. (2002). Data mining: A conceptual overview. In *Communications of the Association for Information Systems* (Vol. 8, pp. 267–296).

10. Kahane, Y., Levin, N., Meiri, R., & Zahavi, J. (2007). Applying Data Mining Technology for Insurance Rate Making: An Example of Automobile Insurance. *Asia-Pacific Journal of Risk and Insurance*, *2*(1). https://doi.org/10.2202/2153-3792.1014

11. Kašćelan, V., Kašćelan, L., & Burić, M. N. (2016). A nonparametric data mining approach for risk prediction in car insurance: A case study from the Montenegrin market. *Economic Research-Ekonomska Istrazivanja*, *29*(1), 545–558. https://doi.org/10.1080/1331677X.2016.1175729

12. Kim, Y., & Street, W. N. (2004). An intelligent system for customer targeting: a data mining approach. *Decision Support Systems*, *37*(2), 215–228. https://doi.org/10.1016/S0167-9236(03)00008-3

13. Kulkarni, R. V., & Devale, A. B. (2012). Applications of Data Mining Techniques in Life Insurance. *International Journal of Data Mining & Knowledge Management Process*, *2*(4), 31–40. https://doi.org/10.5121/ijdkp.2012.2404

14. Kuo, R. J., Lin, S. Y., & Shih, C. W. (2007). Mining association rules through integration of clustering analysis and ant colony system for health insurance database in Taiwan. *Expert Systems with Applications*, *33*(3), 794–808. https://doi.org/10.1016/j.eswa.2006.08.035

15. Liao, S. H., Chen, Y. N., & Tseng, Y. Y. (2009). Mining demand chain knowledge of life insurance market for new product development. *Expert Systems with Applications*, *36*(5), 9422–9437. https://doi.org/10.1016/j.eswa.2008.12.053

16. Lin, Y.-J., Huang, C.-S., & Lin, C.-C. (2008). Determination of Insurance Policy Using Neural Networks and Simplified Models with

Programming Approach to Optimise Insurance Premium Pricing within a Data Mining Framework. *The Journal of the Operational Research Society*, *53*(11), 1197–1203. https://doi.org/10.1057/palgravejors.2601413

insurance claim patterns using data mining: A case study. *Journal of the Operational Research Society*, *51*(5), 532–541. https://doi.org/10.1057/palgrave.jors.2600941

24. Song, H., & Qiuhong, H. (2018). Application of Data Mining Technology in the Loss of Customers in Automobile Insurance Enterprises. *International Journal of Data Science and Analysis*, *4*(1), 1–5. https://doi.org/10.11648/j.ijdsa.20180401.11

25. Sumathi, S., & Sivanandam, S. N. (2006). *Introduction to Data Mining and its Applications* (Vol. 29). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-34351-6

26. Wang, S. L., Pai, H. T., Wu, M. F., Wu, F., & Li, C. L. (2017). The evaluation of trustworthiness to identify health insurance fraud in dentistry. *Artificial Intelligence in Medicine*, *75*, 40–50. https://doi.org/10.1016/j.artmed.2016.12.002

27. Wanke, P., & Barros, C. P. (2016). Efficiency drivers in Brazilian insurance: A two-stage DEA meta frontier-data mining approach. *Economic Modelling*, *53*, 8–22. https://doi.org/10.1016/j.econmod.2015.11.005

28. Watson, H. J., & Wixom, B. H. (2007). The current state of business intelligence. *IEEE Computer Society*, *40*(9), 96–99. https://doi.org/10.1109/MC.2007.331

29. Wu, C. H., Kao, S. C., Su, Y. Y., & Wu, C. C. (2005). Targeting customers via discovery knowledge for the insurance industry. *Expert Systems with Applications*, *29*(2), 291–299. https://doi.org/10.1016/j.eswa.2005.04.002

30. Wu, K.-S., Wang, F.-K., & Shyng, J.-Y. (2013). Applied data mining techniques in insurance company : A comparative study of rough sets and decision tree. *African Journal of Business Management*, *7*(24), 2309–2318. https://doi.org/10.5897/AJBM10.1465

31. Yeo, A. C. A. C., Smith, K. A., Willis, R. J., Brooks, M., Yeo, A. C. A. C., Smithl, K. A., … Brooks, M. (2002). A Mathematical